Infrared Moving Small Target Detection Based on Consistency of Sparse Trajectory

Mo Wu, Xiubin Yang, Zongqiang Fu, Haoyang He, Jiamin Du, Tingting Xu, and Ziming Tu

***Abstract*— Infrared search and track (IRST) systems require reliable detection of small targets in complex backgrounds. Outlier-based methods are prone to high false positive rates due to the resemblance of point-like background features to small targets. The difference image-based method is an effective approach for suppressing point-like background interference; however, it has limitations in detecting slow-moving targets. In this letter, a novel sparse trajectory is proposed for moving target detection in IR videos. With a trajectory growing strategy, two kinds of trajectories from difference images, namely, short sparse trajectories and long sparse trajectories, are correlated to avoid the slow-moving targets being dismissed. The strategy matches the trajectories based on the sparse trajectory intensity composed of similarity measures and optical flow consistency. Finally, real targets are extracted from candidate trajectories using trajectory filtering. Experimental results show that, in the scene with point-like background features, our method achieves the best detection rate and lowest false alarm compared to the state-of-the-art methods.**

***Index Terms*— Infrared moving small target, optical flow consistency, similarity measure, sparse trajectory, trajectory growth.**

I. Introduction

INFRARED moving target detection technology has played an important role in early warning, precise guidance, and other fields. However, detecting infrared small targets is still a challenging task due to their low signal-to-clutter ratio (SCR) and the presence of target-like disturbances in the background, such as building spots and detector noise.

Existing small target detection methods can be classified into two categories: single-frame and sequential detection methods. Single-frame detection methods rely on the human visual system (HVS) theory and typically use different feature descriptors to describe the dissimilarity between the target and surrounding background [1]. Although these methods are straightforward to implement, they are limited by the contrast between target and background and do not perform well at low SCR. To address this issue, some methods attempted to fuse local and global features [2]. Based on the spectral graph theory, Xia et al. [3] proposed modified random walks, which leverage spectral graph theory to describe the sparsity of the target. Huang et al. [4] introduced density peaks searching (DPS), which can detect sparsely distributed density peaks in an image. However, these methods are still susceptible to false detections when objects in the background are highly similar to small targets, as they only use spatial information.

Popular existing sequential methods treat multiple frames in a sequence as a whole, introducing the local contrast space into 3-D. Related works include spatial-temporal local contrast filter (STLCF) [5], spatial–temporal local difference measure (STLDM) [6], and novel spatiotemporal saliency method (NSTSM) [7]. Although the mentioned methods exploit spatial–temporal correlations in different ways, outlier-based methods still face high false positive rates when disturbed by point-like background features [8]. The time-domain difference (TDD) method [9] can effectively suppress point-like background interference, but its limitations are obvious. Slow-moving target will be suppressed in the difference image, resulting in missed detection. One target produces two spots in a difference image, resulting in false detections [10].

We propose a novel sequential method for detecting moving infrared targets, which addresses the limitations of TDD. Unlike previous methods that directly detect Gaussian distribution patterns, our proposed method identifies adjacent positive and negative peak pairs (P-NDPs) in the registered difference images to construct short and long sparse trajectories. Considering that using only short trajectories may miss slowly moving targets, while using only long trajectories may result in false alarms due to sensitivity to noise, we incorporate homography transformation based on motion continuity to correlate sparse trajectories across frames, thereby preventing the suppression of slow-moving targets.

Our method is based on three assumptions about small targets: 1) the imaging distance is far; 2) the target structure is approximately a Gaussian distribution; and 3) the target is rare in the whole image with slow and continuous movement relative to the background [11]. Because of the first assumption, the target and the background can be regarded as lying on the same plane in the distance, which ensures that images can be registered using homography transformation. The latter two assumptions are used to extract sparse trajectories and describe the optical flow consistency of the trajectories.

II. Proposed Method

A. Overview

The flowchart of the proposed method is shown in Fig. 1. Based on the sparsity of the target, DPS introducing local contrast (DPS-LC) is used to search for P-NDPs, which constitute the sparse trajectories. Subsequently, the similarity measure and the optical flow consistency of the trajectory are fused to describe the sparse trajectory intensity, which is used for trajectory growth. Then, the real target is extracted by trajectory filtering.

Manuscript received 15 December 2022; revised 1 March 2023; accepted 11 March 2023. Date of publication 16 March 2023; date of current version 24 March 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 62171430 and Grant 62101071, and in part by the Natural Science Foundation of Jilin Province under Grant 20220101031JC. *(Corresponding author: Xiubin Yang.)*

Mo Wu, Xiubin Yang, Zongqiang Fu, Jiamin Du, Tingting Xu, and Ziming Tu are with the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China, and also with the Daheng College, University of Chinese Academy of Sciences, Beijing 100039, China (e-mail: yangxiubin@ciomp.ac.cn).

Haoyang He is with the School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China.

Digital Object Identifier 10.1109/LGRS.2023.3257850

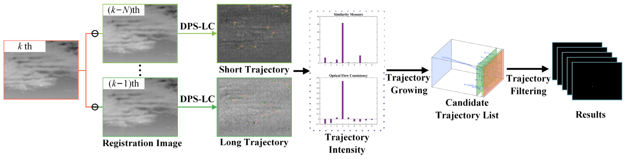


Fig. 1. Flowchart of the proposed detection method.

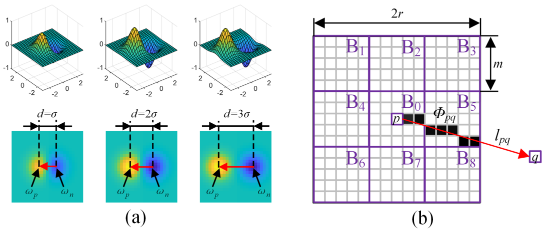


Fig. 2. (a) Difference model of the target whose moving distance is σ, 2σ, and 3σ. (b) Local structure for calculating density-distance for p.

B. Extract Sparse Trajectory

1) Time Domain Difference: Ideally, the difference images contain the target component only. According to previous research, the small target can be modeled as a 2-D Gaussian function. The difference model of moving target can be expressed as

$$\mathrm{diff}_{ij} = g(x, y, x_i, y_i) - g(x, y, x_j, y_j) \tag{1}$$

$$g(x, y, x_i, y_i) = A\exp\left(-(x - x_i)^2/2\sigma_x^2 - (y - y_i)^2/2\sigma_y^2\right) \tag{2}$$

where g is the Gaussian function, A is the peak value of the target area, (x_i, y_i) is the center position of the target in the ith frame, and σ_x and σ_y are the horizontal and vertical parameters of the target, respectively. Fig. 2(a) shows the 3-D difference model of the moving target, including two adjacent peaks: the positive peak ω_p in convex region and the negative peak ω_n in depression area. If the moving distance of the target is larger than σ, more than 50% of the energy will be retained in each peak. We can extract the trajectory of target by using the high-intensity P-NDPs, which are rare in the background.

Taking the kth frame as an example, homography transform is used to warp f_k to the perspective of f_{k-1}. The warped image is denoted as $f_{k,k-1}$. To prevent the slow-moving target from being suppressed, we obtain difference images $\mathrm{diff}_{k,k-1}$ and $\mathrm{diff}_{k,k-N}$ simultaneously for detecting sparse trajectories, as shown in the following equations, which ensures that the motion target with speed larger than σ/N (pix/frame) can be detected:

$$\mathrm{diff}_{k,k-1} = f_{k,k-1} - f_{k-1} \tag{3}$$

$$\mathrm{diff}_{k,k-N} = f_{k,k-N} - f_{k-N}. \tag{4}$$

2) Density Peaks Searching Introducing Local Contrast: Based on the sparsity of the small target, we use DPS [4] to search for positive and negative peaks in difference image. To avoid the target being suppressed by nearby high-brightness clutter, the local contrast is introduced into DPS. The density distance of pixel p is redefined as

$$\delta_p = d_{pq} + \sum_{t\in\Phi_{pq}} (\rho_p - \rho_t)/r \tag{5}$$

$$q = \arg\min_{q'}(d_{pq'}) \tag{6}$$

where $d(p, q)$ represents the Euclidean distance between p and q, ρ represents the density, and Φ_{pq} is the intersection of the line l_{pq} between p and q and the local window, as shown in Fig. 2(b). The density peak index of each pixel is denoted by γ. To obtain as many outliers as possible in ρ–δ space, the k-nearest neighbor distance ($d_{k-\mathrm{NN}}$) is used to describe the outlier degree of each pixel

$$d_{k\text{-NN}}(p) = \sum_{q\in k\text{-NN}(p)} \left|\gamma_p - \gamma_q\right|. \tag{7}$$

Finally, all peaks obtained are expressed as

$$\Omega = \left\{(x_p, y_p)|d_{k\text{-NN}}(p) > \mathrm{std}(d_{k\text{-NN}})\right\}. \tag{8}$$

3) Sparse Trajectory: DPS-LC is applied to $\mathrm{diff}_{k,k-1}$ and $-\mathrm{diff}_{k,k-1}$ to obtain positive peaks Ω_{pos} and negative peaks Ω_{neg}. Positive and negative density peaks of which distance are smaller than the maximum optical flow length constitute short sparse trajectories between f_{k-1} and f_k. Specifically, if a density peak pair satisfies: $\omega_i \in \Omega_{\mathrm{neg}}$, $\omega_j \in \Omega_{\mathrm{pos}}$, and $\mathrm{d}(\omega_i, \omega_j) < \max(|\Gamma_{k-1,k}|)$, the short trajectory T_s they constitute is

$$T_s = \left\{\varpi^{k-1}, \varpi^k\right\} \tag{9}$$

where $\boldsymbol{\Gamma}$ is the Farneback optical flow, $\varpi^{k-1} = \omega_i$ and $\varpi^k = H_{k-1,k}\omega_j$, and $H_{k-1,k}$ is the homography matrix from f_{k-1} to f_k. Similarly, the P-NDPs in $\mathrm{diff}_{k,k-N}$ form long sparse trajectories T_l. The track points inside T_l are obtained by interpolation

$$T_l = \left\{\varpi^{k-N}, \ldots, \varpi^{k-1}, \varpi^k\right\} \tag{10}$$

$$\varpi^{k-n} = H_{k-N,k-n}\left(\omega_i + (N - n)\left(\omega_j - \omega_i\right)/N\right),\ 0 \le n \le N. \tag{11}$$

We refer to the endpoints of sparse trajectories as "real points," and the internal points obtained by interpolation as "imaginary points." The type of trajectory point ϖ is denoted by the symbol $P(\varpi)$

$$P(\varpi) = \begin{cases} 1, & \text{if } \varpi \text{ is a real point} \\ 0, & \text{else.} \end{cases} \tag{12}$$

C. Sparse Trajectory Intensity

Sparse trajectory intensity is used to describe the likelihood that a sparse trajectory is a real moving target, which combines similarity measures and optical flow consistency.

1) Similarity Measure: For a sparse trajectory T composed of a pair of density peaks (ω_i, ω_j), the local areas of size $m \times m$, which have center on ω_i and ω_j, are denoted by B_0^i and B_0^j, respectively. Referring to the model in Fig. 2(a), the convex area B_0^i and the concave area B_0^j formed by the same target must be similar. In addition, both B_0^i and B_0^j are quite different from the adjacent subblocks B_{1-8}^i.

Wasserstein distance is used to describe the similarity of central blocks. For two distributions μ and ν, their Wasserstein distance is

$$W(\mu, \nu) = \inf_{\chi\in\Pi(\mu,\nu)} \iint \chi(x, y)d(x, y)dxdy \tag{13}$$

where χ is the joint distribution of μ and ν. We divide the pixels in B_0^i and B_0^j into m^2 gray-scale intervals. Their gray-scale distribution is expressed as μ^i and ν^j. The similarity S^{ij} of B_0^i and B_0^j is defined as follows:

$$S^{ij} = \exp\left(-W\left(\mu^i, \nu^j\right)\right). \quad (14)$$

Structural dissimilarity and gradient flux [12] are used to indicate the difference between B_0 and peripheral subblocks. The structural dissimilarity between the peripheral subblock B_t^i and the central block B_0^i is

$$D^i = \min_t\left(\left(\mathrm{std}\left(e_{t0}^i\right) - \mathrm{std}\left(e_0^i\right)\right)/\mathrm{std}\left(e_{t0}^i\right)\right) \quad (15)$$

where e_0^i and e_{t0}^i represent the vectorized B_0^i and $[B_0^i, B_t^i]$. Gradient flux is used to evaluate the number of vectors flowing into or out of the closed area, and the flux in the convex area is often a small negative value. To obtain the gradient flux of B_0, the 2-D Gaussian function (2) is used to fit B_0. Denoting the least squares estimates of parameters in (2) with ($\hat{A}$, $\hat{x}$, $\hat{y}$, $\hat{\sigma}_x$, and $\hat{\sigma}_y$), the fitted surface without offset is denoted as

$$\hat{g}(x, y) = \hat{A}\exp\left(-x^2/2\hat{\sigma}_x^2 - y^2/2\hat{\sigma}_y^2\right). \quad (16)$$

The maximum gradient flux of $\hat{g}$ is taken as the gradient flux of B_0, as shown in the following equation:

$$F = \max_C \oint_C |\mathbf{F}\cdot\mathbf{n}|ds, \text{ where } C{:}z = \hat{g}(x, y), z > 0 \quad (17)$$

where F can be obtained by

$$F = 2\pi\hat{A}\left(1 + r_C^2\right)/(r_C e) \quad (18)$$

where $r_c = \hat{\sigma}_x/\hat{\sigma}_y$ is the aspect ratio of C. The local structural difference is defined as

$$D^{ij} = \min\left(D^i, D^j\right) \times \min\left(F^i, F^j\right). \quad (19)$$

The similarity measure is defined as the product of the similarity of central blocks and local structural difference, namely,

$$\mathrm{SM} = S^{ij} \times D^{ij}. \quad (20)$$

2) Optical Flow Consistency: Trajectories containing real objects are usually consistent with optical flow. The optical flow consistency of a trajectory can be expressed as

$$\mathrm{FC} = v^{ij}\cdot\Gamma(\omega_i)/\max\left(\left|v^{ij}\right|, |\Gamma(\omega_i)|\right)^2 \quad (21)$$

where v^{ij} represents the vector from ω_j to ω_i, and $\Gamma(\omega_i)$ represents the optical flow of ω_i.

Finally, the sparse trajectory intensity is expressed as

$$I_{st} = (\mathrm{FC} + 1) \times \mathrm{SM}/2. \quad (22)$$

From (22), it can be seen that trajectories with larger I_{st} are more likely to be formed by real targets. The intensity of a trajectory point ϖ in T is denoted by the symbol $\mathrm{I}(\varpi)$, which is equal to the intensity of T.

D. Trajectory Growing Strategy

To obtain the complete trajectory of the target, a trajectory-growing strategy is introduced. Let $\mathbf{\Lambda}^k = \{\boldsymbol{T}^{k_1}, \ldots, T_n^k\}$ denote the sparse trajectory list composed of all sparse trajectories extracted from f_k, and let $\boldsymbol{L}$ denote the list of candidate trajectories. When $k = 2$, let $\boldsymbol{L} = \Lambda^k$. When $k > 2$, the sparse trajectories in Λ^k are matched with candidate trajectories in L. Let $T_i^k = \{\varpi^{k-N(i)}, \ldots, \varpi^k\}$ denote be a trajectory in Λ^k, and $L_j = \{l^{k-N(j)}, \ldots, l^{k-1)}\}$ denote a trajectory in $\boldsymbol{L}$.

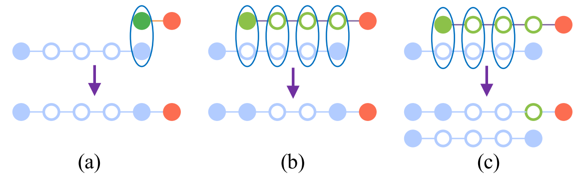


Fig. 3. Trajectory growing strategy. (a) Short trajectory is matched with the candidate trajectory. (b) Long trajectory is fully matched with the candidate trajectory. (c) Long trajectory is partially matched with the candidate trajectory. Real and imaginary points are represented as solid and hollow circles, respectively. The trajectory points corresponding to Ω_{pos} and Ω_{neg} are marked with orange and green, respectively. The matching trajectory points are marked with blue ellipses.

The trajectory matching rule is: if there exists $a \in [1, N(i)]$, such that for any $b \in [a, N(i)]$, $\mathrm{d}(\varpi^{k-b}, l^{k-b}) < R$, T_i^k and L_j are matched, then, ϖ^{k-b} and l^{k-b} are called matching points. The position of the last trajectory point that matches T_i^k and L_j is denoted by $\alpha = \min(\mathrm{a})$. $\alpha = 1$ means that T_i^k and L_j are fully matched, and L_j grows as $L_j = \{l^{k-N(j)}, \ldots, l^{k-1)}, \varpi^k\}$. $\alpha > 1$ means that T_i^k and L_j are partially matched. L_j is split into two trajectories, one of which is the trajectory $L_j = \{l^{k-N(j)}, \ldots, l^{k-\alpha}, \varpi^{k-\alpha+1}, \ldots, \varpi^k\}$ grown to the kth frame, and the other is the original trajectory $L_j = \{l^{k-}N(j), \ldots, l^{k-1)}\}$. Then, update the type and intensity of the trajectory point in L_j. For matching points ϖ^{k-b} and l^{k-b}, let $\mathrm{I}(l^{k-b}) = \max[\mathrm{I}(\varpi^{k-b}), \mathrm{I}(l^{k-b})]$ and $P(l^{k-b}) = P(\varpi^{k-b})|P(l^{k-b})$. Specifically, Fig. 3 shows three trajectory growth examples. As shown in Fig. 3(a), if the candidate trajectory matches a short trajectory, the trajectory point of the kth frame of the short trajectory is added to L_j. $\mathrm{I}(l^{k-1}) = \max[\mathrm{I}(\varpi^{k-1}), \mathrm{I}(l^{k-1})]$ and $P(l^{k-1}) = P(\varpi^{k-1})|P(l^{k-1})$. If the candidate trajectory matches the long trajectory, there are two cases depending on the value of α. As shown in Fig. 3(b), if $\alpha = 1$, the candidate trajectory exactly matches the long trajectory. Since the endpoint $\varpi^{k-N(i)}$ of the long trajectory is a real point, it also becomes a real point, that is, let $P(l^{k-N(i)}) = P(\varpi^{k-N(i)})|P(l^{k-N(i)})$. When $\alpha > 1$, as shown in Fig. 3(c). The candidate trajectory is split into two trajectories. For the matching points in the grown trajectory L_j, let $\mathrm{I}(l^{k-b}) = \max[\mathrm{I}(\varpi^{k-b}), \mathrm{I}(l^{k-b})]$ and $P(l^{k-b}) = P(\varpi^{k-b})|P(l^{k-b})$.

E. Trajectory Filtering

Considering a candidate trajectory $L_j = \{l^{k-N(j)}, \ldots, l^k\}$ in $\boldsymbol{L}$, its average intensity are defined as

$$M = \frac{1}{N(j)+1}\sum_{i=0}^{N(j)} \mathrm{I}(\varpi) \times P(\varpi). \quad (23)$$

The average intensity of the last five trajectory points of L_j is denoted with $\tilde{M}$, and the proportion of real points in L_j is denoted with η. The following conditions are used to select the true trajectory:

$$\begin{cases} M > \mathrm{Th} \\ \tilde{M} > \mathrm{Th} \\ \eta > 80\% \\ N(j) \geq 2N \end{cases} \quad (24)$$

where Th is a fixed threshold. The trajectory point of the selected real trajectory in f_k is the target detected in this frame.

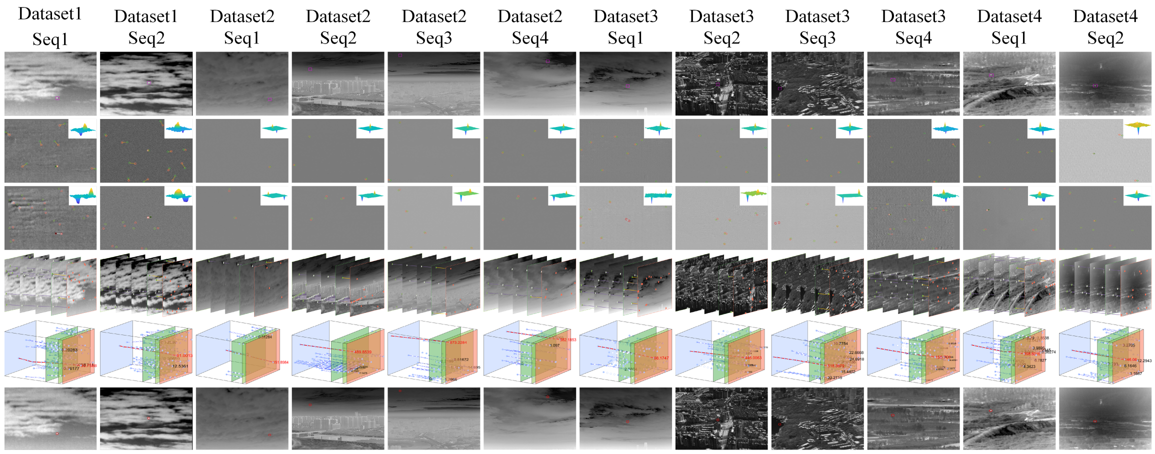


Fig. 4. Detection results of our method in the 15th frame of all sequences. First row: original images. Second and third rows: difference images $\mathrm{diff}_{k,k-1}$ and $\mathrm{diff}_{k,k-N}$. Positive and negative peaks are marked with orange and green circles, respectively. Missed trajectories are marked with white lines. Fourth row: sparse trajectories. Fifth row: candidate trajectories. Sixth row: detection results.

TABLE I

DETAILS OF FOUR IR TARGET DATASETS

	Sequences	Frames	Resolution	Speed	Size	Background	Mean SCR
Dataset 1	2	131	320 × 240	> 3 pix/frame	8 × 6	changing	12.1626
Dataset 2	4	242	640 × 512	> 3 pix/frame	3 × 3	changeless	55.5701
Dataset 3	4	187	640 × 512	> 3 pix/frame	3 × 3	changeless	37.9729
Dataset 4	2	1400	256 × 256	0-3 pix/frame	3 × 3	changeless	22.1341

III. EXPERIMENTAL RESULTS

We conduct comparative experiments on four datasets to evaluate the effectiveness and robustness of our method. The details of the datasets are recorded in Table I. Each dataset contains multiple sequences with the same target features. All the datasets used meet the three assumptions about small targets mentioned in the introduction. All experiments were performed using MATLAB 2020a on a computer with an Intel i5 CPU and 16 GB of memory. The receiver operating characteristic curve (ROC) and area under the curve (AUC) are used to evaluate the detection performance of methods.

According to the three-layer window theory, the core layer of a target is smaller than 5 × 5, so the size of subblock B in Fig. 2(b) is set to 5 × 5 to cover the core part of the target. In the following experiments, the forward frame number N for extracting long sparse trajectories is set to 4, R is set to 5, and Th is set to 50.

Fig. 4 presents the detection results of our method on the 15th frame of all sequences. The second and third rows show the difference images $\mathrm{diff}_{k,k-1}$ and $\mathrm{diff}_{k,k-N}$, respectively. It can be seen that the background in the difference image is well suppressed. The 3-D images of local area near the target are given in the upper right corner of the difference images. P-NDPs formed by moving targets are prominent on the difference image, which means that our method of extracting sparse trajectories is reasonable. For sequences containing fast-moving targets, such as S1 in Dataset1 and S1 in Dataset3, the positive and negative peaks in $\mathrm{diff}_{k,k-N}$ are further apart than those in $\mathrm{diff}_{k,k-1}$. The long sparse trajectories formed by the target are not extracted, but the short sparse trajectories are successfully extracted in $\mathrm{diff}_{k,k-1}$, which is in line with our original design intention. In addition, it can be seen in the fourth row that although the short and long sparse trajectories are not extracted simultaneously, the full trajectory of the target can be obtained by our trajectory growing strategy, and its average trajectory intensity is much higher than that of the false trajectory.

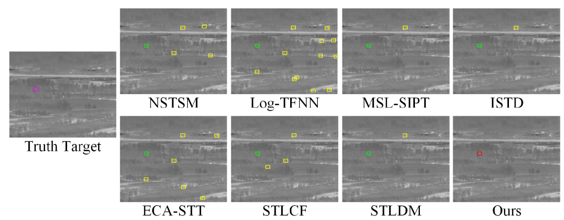


Fig. 5. Examples for results comparison for a scene with point-like background features. Missing and false detections are marked with green and yellow square boxes, respectively.

The baseline methods we compare include NSTSM [7], tensor fibered nuclear norm based on the Log operator (Log-TFNN) [13], multiple subspace learning and spatial-temporal patch-tensor (MSL-SIPT) [14], infrared small-target detector (ISTD) [15], edge and corner awareness-based spatial–temporal tensor (ECA-STT) [16], STLCF [5], and STLDM [6]. Log-TFNN is a single-frame detection method, and others are the sequential detection methods. Fig. 5 shows a comparison of detection results in complex scenes from Dataset 3. Due to the presence of multiple building spots and objects similar to small targets in the scene, the baseline methods all suffer from false detections, ignoring the dim real target. However, our method can detect the target accurately. To better demonstrate the advantages of our method, we compared the detection results of different methods on sequence 22 of the ISD dataset [17] containing a slow-moving target. As shown in Fig. 6, our method successfully detected all slow-moving targets that were missed by other methods. Fig. 7 shows the ROC curves using eight methods. It can be seen that our method achieves the best detection performance on Dataset 1, Dataset 3, and Dataset 4 compared with baseline method. Most of the sequential methods achieve good results on Dataset 2 due to constant background and the higher SCR of the scene. The AUC value of our method is only 0.0001 lower than the best-performing MSL-SIPT. However, only our method performs well in complex scene of Dataset 3, with R_d greater than 0.99 when R_f is low. In addition, most sequential methods

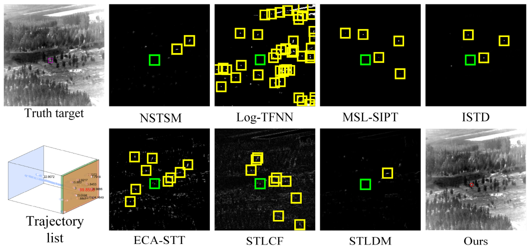
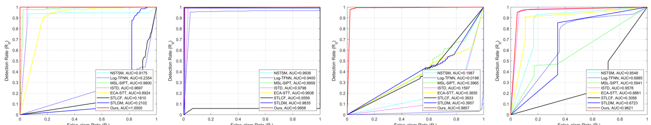


Fig. 6. Detection results for a sequence containing a slow-moving target.



(a) (b) (c) (d)

Fig. 7. (a)–(d) ROC curves of eight methods in datasets 1–4.

TABLE II

AVERAGE COMPUTATION TIME PER FRAME (SECONDS)

	NSTSM	Log-TFNN	MSL-SIPT	ISTD	ECA-STT	STLCF	STLDM	Ours	Ours on GPU
Dataset 1	0.9718	1.0255	9.2784	0.1121	6.9870	0.1903	1.4661	0.5659	0.0936
Dataset 2	4.3062	4.1904	35.110	0.4217	34.924	0.8637	6.3553	1.7032	0.2992
Dataset 3	4.3018	4.3228	31.933	0.3842	40.498	0.8784	6.3600	1.9861	0.3384
Dataset 4	0.8193	1.1001	7.0745	0.1058	6.6259	0.1690	1.2577	0.5790	0.0681

perform better than Log-TFNN, which only utilizes spatial information.

We have presented the running times of all methods in Table II, where ISTD benefits from acceleration by the GTX 1650 GPU. Ours performs slightly slower than ISTD and STLCF. This is because our method uses two DPS to extract sparse trajectories, and calculating density distances for all pixels in the entire image is time-intensive. However, we have leveraged GPU acceleration to speed up this process, allowing our method to outperform all other methods in terms of speed.

IV. CONCLUSION

In this letter, an infrared small target detection method based on the consistency of sparse trajectories is proposed. The main idea is to detect sparse trajectories of target on registered interframe difference images to exploit temporal contextual information. To prevent slow-moving targets from being suppressed in the difference image, both long and short sparse trajectories are used for trajectory growth. Experiments show that this strategy is reasonable, and our method can effectively detect targets of different sizes and speeds. Compared with the existing single-frame and sequential methods, our method has better detection ability in changing complex backgrounds. In future work, the direction consistency of sparse trajectories can be used to describe the trajectory intensity, which can improve the robustness to temporal noise.

REFERENCES

[1] C. L. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, "A local contrast method for small infrared target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 574–581, Jan. 2013, doi: 10.1109/TGRS.2013.2242477.

[2] J. Xue, D. He, M. Liu, and Q. Shi, "Dual network structure with interweaved global-local feature hierarchy for transformer-based object detection in remote sensing image," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 6856–6866, 2022, doi: 10.1109/JSTARS.2022.3198577.

[3] C. Xia, X. Li, and L. Zhao, "Infrared small target detection via modified random walks," *Remote Sens.*, vol. 10, no. 12, p. 2004, Dec. 2018, doi: 10.3390/rs10122004.

[4] S. Huang, Z. Peng, Z. Wang, X. Wang, and M. Li, "Infrared small target detection by density peaks searching and maximum-gray region growing," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 12, pp. 1919–1923, Dec. 2019, doi: 10.1109/LGRS.2019.2912989.

[5] L. Z. Deng, H. Zhu, C. Tao, and Y. T. Wei, "Infrared moving point target detection based on spatial–temporal local contrast filter," *Infr. Phys. Technol.*, vol. 76, pp. 168–173, May 2016, doi: 10.1016/j.infrared.2016.02.010.

[6] P. Du and A. Hamdulla, "Infrared moving small-target detection using spatial–temporal local difference measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 10, pp. 1817–1821, Oct. 2020, doi: 10.1109/LGRS.2019.2954715.

[7] D. Pang, T. Shan, P. Ma, W. Li, S. Liu, and R. Tao, "A novel spatiotemporal saliency method for low-altitude slow small infrared target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2020.3048199.

[8] D. Li, B. Mo, and J. Zhou, "Boost infrared moving aircraft detection performance by using fast homography estimation and dual input object detection network," *Infr. Phys. Technol.*, vol. 123, Jun. 2022, Art. no. 104182, doi: 10.1016/j.infrared.2022.104182.

[9] S. S. Sengar and S. Mukhopadhyay, "Moving object detection based on frame difference and W4," *Signal, Image Video Process.*, vol. 11, no. 7, pp. 1357–1364, Oct. 2017, doi: 10.1007/s11760-017-1093-8.

[10] B. Zhao, S. Xiao, H. Lu, and D. Wu, "Spatial–temporal local contrast for moving point target detection in space-based infrared imaging system," *Infr. Phys. Technol.*, vol. 95, pp. 53–60, Dec. 2018, doi: 10.1016/j.infrared.2018.10.011.

[11] H. Zhang, D. Ye, Y. Xiao, and Z. Sun, "Adaptive control on SE(3) for spacecraft pose tracking with harmonic disturbance and input saturation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 5, pp. 4578–4594, Oct. 2022, doi: 10.1109/TAES.2022.3165524.

[12] B. Xiong, X. Huang, and M. Wang, "Local gradient field feature contrast measure for infrared small target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 3, pp. 553–557, Mar. 2021, doi: 10.1109/LGRS.2020.2976208.

[13] X. Kong, C. Yang, S. Cao, C. Li, and Z. Peng, "Infrared small target detection via nonconvex tensor fibered rank approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5000321, doi: 10.1109/TGRS.2021.3068465.

[14] Y. Sun, J. Yang, and W. An, "Infrared dim and small target detection via multiple subspace learning and spatial–temporal patch-tensor model," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 3737–3752, May 2020, doi: 10.1109/TGRS.2020.3022069.

[15] L. Xu, Y. Wei, H. Zhang, and S. Shang, "Robust and fast infrared small target detection based on Pareto frontier optimization," *Infr. Phys. Technol.*, vol. 123, Jun. 2022, Art. no. 104192, doi: 10.1016/j.infrared.2022.104192.

[16] P. Zhang, L. Zhang, X. Wang, F. Shen, T. Pu, and C. Fei, "Edge and corner awareness-based spatial–temporal tensor model for infrared small-target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 10708–10724, Dec. 2021, doi: 10.1109/TGRS.2020.3037938.

[17] H. Bingwei, "A dataset for infrared image dim-small aircraft target detection and tracking under ground/air background," *Sci. Data Bank*, vol. 5, p. 12, Oct. 2019, doi: 10.11922/sciencedb.902.